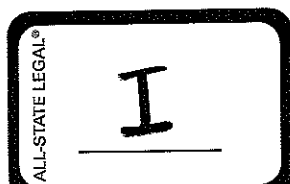


**OFFICE OF MANAGEMENT AND BUDGET**  
**STANDARDS AND GUIDELINES FOR STATISTICAL SURVEYS**

**September 2006**

**Table of Contents**

|   |           |
|---|-----------|
| <b>LIST OF STANDARDS FOR STATISTICAL SURVEYS .....</b>                      | <b>i</b>  |
| <b>INTRODUCTION.....</b>  | <b>1</b>  |
| <b>SECTION 1 DEVELOPMENT OF CONCEPTS, METHODS, AND DESIGN .....</b>         | <b>5</b>  |
| Section 1.1 Survey Planning .....   | 5         |
| Section 1.2 Survey Design.....  | 7         |
| Section 1.3 Survey Response Rates .....                                     | 8         |
| Section 1.4 Pretesting Survey Systems .....                                 | 9         |
| <b>SECTION 2 COLLECTION OF DATA.....</b>                                    | <b>9</b>  |
| Section 2.1 Developing Sampling Frames.....                                 | 9         |
| Section 2.2 Required Notifications to Potential Survey Respondents .....    | 10        |
| Section 2.3 Data Collection Methodology.....                                | 11        |
| <b>SECTION 3 PROCESSING AND EDITING OF DATA.....</b>                        | <b>13</b> |
| Section 3.1 Data Editing .....  | 13        |
| Section 3.2 Nonresponse Analysis and Response Rate Calculation.....         | 14        |
| Section 3.3 Coding.....   | 17        |
| Section 3.4 Data Protection.....  | 18        |
| Section 3.5 Evaluation .....  | 19        |
| <b>SECTION 4 PRODUCTION OF ESTIMATES AND PROJECTIONS .....</b>              | <b>20</b> |
| Section 4.1 Developing Estimates and Projections.....                       | 20        |
| <b>SECTION 5 DATA ANALYSIS .....</b>  | <b>21</b> |
| Section 5.1 Analysis and Report Planning.....                               | 21        |
| Section 5.2 Inference and Comparisons.....                                  | 22        |
| <b>SECTION 6 REVIEW PROCEDURES .....</b>                                    | <b>23</b> |
| Section 6.1 Review of Information Products .....                            | 23        |
| <b>SECTION 7 DISSEMINATION OF INFORMATION PRODUCTS .....</b>                | <b>24</b> |
| Section 7.1 Releasing Information.....                                      | 24        |
| Section 7.2 Data Protection and Disclosure Avoidance for Dissemination..... | 25        |
| Section 7.3 Survey Documentation.....                                       | 26        |
| Section 7.4 Documentation and Release of Public-Use Microdata .....         | 27        |
| <b>APPENDIX DEFINITIONS OF KEY TERMS .....</b>                              | <b>29</b> |



DesVosges-Rausser009295

data collections; methods of collection for achieving acceptable response rates; training of enumerators and persons coding and editing the data; and cost estimates, including the costs of pretests, nonresponse follow-up, and evaluation studies.

**Guideline 1.2.6:** Whenever possible, construct an estimate of total mean square error in approximate terms, and evaluate accuracy of survey estimates by comparing with other information sources. If probability sampling is used, estimate sampling error; if nonprobability sampling is used, calculate the estimation error.

**Guideline 1.2.7:** When possible, estimate the effects of potential nonsampling errors including measurement errors due to interviewers, respondents, instruments, and mode; nonresponse error; coverage error; and processing error.

### **Section 1.3 Survey Response Rates**

**Standard 1.3:** Agencies must design the survey to achieve the highest practical rates of response, commensurate with the importance of survey uses, respondent burden, and data collection costs, to ensure that survey results are representative of the target population so that they can be used with confidence to inform decisions. Nonresponse bias analyses must be conducted when unit or item response rates or other factors suggest the potential for bias to occur.

**Key Terms:** cross-sectional, key variables, longitudinal, nonresponse bias, response rates, stage of data collection, substitution, target population, universe

---

The following guidelines represent best practices that may be useful in fulfilling the goals of the standard:

**Guideline 1.3.1:** Calculate sample survey unit response rates without substitutions.

**Guideline 1.3.2:** Design data collections that will be used for sample frames for other surveys (e.g., the Decennial Census, and the Common Core of Data collection by the National Center for Education Statistics) to meet a target unit response rate of at least 95 percent, or provide a justification for a lower anticipated rate (See Section 2.1.3).

**Guideline 1.3.3:** Prior to data collection, identify expected unit response rates at each stage of data collection, based on content, use, mode, and type of survey.

**Guideline 1.3.4:** Plan for a nonresponse bias analysis if the expected unit response rate is below 80 percent (see Section 3.2.9).

**Guideline 1.3.5:** Plan for a nonresponse bias analysis if the expected item response rate is below 70 percent for any items used in a report (see Section 3.2.9).

$K$  = the last stage of data collection used in the analysis;  
 $I^L$  = the number of responding cases common to all waves in the analysis  
 $R_k^1$  = Refusals at wave 1 at stage  $k$   
 so that  $I_k^1 + R_k^1 + O_k^1 + NC_k^1 + e_k(U_k^1)$  is the entire sample entered at wave 1

**Guideline 3.2.6:** Calculate item response rates (RRI) as the ratio of the number of respondents for whom an in-scope response was obtained ( $I^x$  for item  $x$ ) to the number of respondents who were asked to answer that item. The number asked to answer an item is the number of unit-level respondents ( $I$ ) minus the number of respondents with a valid skip for item  $x$  ( $V^x$ ). When an abbreviated questionnaire is used to convert refusals, the eliminated questions are treated as item nonresponse:

$$RRI^x = \frac{I^x}{I - V^x}$$

**Guideline 3.2.7:** Calculate the total item response rates ( $RRT^x$ ) for specific items as the product of the overall unit response rate (RRO) and the item response rate for item  $x$  ( $RRI^x$ ):

$$RRT^x = RRO * RRI^x$$

**Guideline 3.2.8:** When calculating a response rate with supplemented samples, base the reported response rates on the original and the added sample cases. However, when calculating response rates where the sample was supplemented during the initial sample selection (e.g., using matched pairs), calculate unit response rates without the substituted cases included (i.e., only the original cases are used).

**Guideline 3.2.9:** Given a survey with an overall unit response rate of less than 80 percent, conduct an analysis of nonresponse bias using unit response rates as defined above, with an assessment of whether the data are missing completely at random. As noted above, the degree of nonresponse bias is a function of not only the response rate but also how much the respondents and nonrespondents differ on the survey variables of interest. For a sample mean, an estimate of the bias of the sample respondent mean is given by:

$$B(\bar{y}_r) = \bar{y}_r - \bar{y}_t = \left( \frac{n_{nr}}{n} \right) (\bar{y}_r - \bar{y}_{nr})$$

Where:

$\bar{y}_t$  = the mean based on all sample cases;  
 $\bar{y}_r$  = the mean based only on respondent cases;  
 $\bar{y}_{nr}$  = the mean based only on the nonrespondent cases;  
 $n$  = the number of cases in the sample; and  
 $n_{nr}$  = the number of nonrespondent cases.

For a multistage (or wave) survey, focus the nonresponse bias analysis on each stage, with particular attention to the “problem” stages. A variety of methods can be used to examine nonresponse bias, for example, make comparisons between respondents and nonrespondents across subgroups using available sample frame variables. In the analysis of unit nonresponse, consider a multivariate modeling of response using respondent and nonrespondent frame

variables to determine if nonresponse bias exists. Comparison of the respondents to known characteristics of the population from an external source can provide an indication of possible bias, especially if the characteristics in question are related to the survey's key variables.

**Guideline 3.2.10:** If the item response rate is less than 70 percent, conduct an item nonresponse analysis to determine if the data are missing at random at the item level for at least the items in question, in a manner similar to that discussed in Guideline 3.2.9.

**Guideline 3.2.11:** In those cases where the analysis indicates that the data are not missing at random, the amount of potential bias should inform the decision to publish individual items.

**Guideline 3.2.12:** For data collections involving sampling, adjust weights for unit nonresponse, unless unit imputation is done. The unit nonresponse adjustment should be internally consistent, based on theoretical and empirical considerations, appropriate for the analysis, and make use of the most relevant data available.

**Guideline 3.2.13:** Base decisions regarding whether or not to adjust or impute data for item nonresponse on how the data will be used, the assessment of nonresponse bias that is likely to be encountered in the review of collections, prior experience with this collection, and the nonresponse analysis discussed in this section. When used, imputation and adjustment procedures should be internally consistent, based on theoretical and empirical considerations, appropriate for the analysis, and make use of the most relevant data available. If multivariate analysis is anticipated, care should be taken to use imputations that minimize the attenuation of underlying relationships.

**Guideline 3.2.14:** In the case of imputing longitudinal data sets, use cross-wave imputations or cross-sectional imputations.

**Guideline 3.2.15:** Clearly identify all imputed values on a data file (e.g., code them).

For more information on calculating response rates and conducting nonresponse bias analyses, see *FCSM Statistical Policy Working Paper 31, Measuring and Reporting Sources of Error in Surveys*.

### Section 3.3 Coding

**Standard 3.3:** Agencies must add codes to collected data to identify aspects of data quality from the collection (e.g., missing data) in order to allow users to appropriately analyze the data. Codes added to convert information collected as text into a form that permits immediate analysis must use standardized codes, when available, to enhance comparability.

**Key Terms:** coding, quality assurance process

---

The following guidelines represent best practices that may be useful in fulfilling the goals of the standard: